

**Technical Report 1133**

**Web-Based Collaborative Learning: An  
Assessment of a Question-Generation Approach**

**James Belanich  
Robert A. Wisher  
U.S. Army Research Institute**

**Kara L. Orvis  
George Mason University  
Consortium Research Fellows Program**

**January 2003**



**United States Army Research Institute  
for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited.

**20030303 131**

**U.S. Army Research Institute  
for the Behavioral and Social Sciences**

**A Directorate of the U.S. Total Army Personnel Command**

**ZITA M. SIMUTIS  
Director**

---

Technical Review by

J. Douglas Dressel, U.S. Army Research Institute  
Arthur Graesser, University of Memphis  
James J. Dees, TRADOC

**NOTICES**

**DISTRIBUTION:** Primary distribution of this Technical Report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: TAPC-ARI-PO, 5001 Eisenhower Ave., Alexandria, VA 22333-5600.

**FINAL DISPOSITION:** This Technical Report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

**NOTE:** The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

## REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) January 2003	2. REPORT TYPE Interim	3. DATES COVERED (from. . . to) March 2001 – August 2002		
4. TITLE AND SUBTITLE  Web-based collaborative learning: An assessment of a question-generation approach			5a. CONTRACT OR GRANT NUMBER	
			5b. PROGRAM ELEMENT NUMBER 633007	
6. AUTHOR(S)  James Belanich and Robert A. Wisher (U.S. Army Research Institute); Kara L. Orvis (George Mason University)			5c. PROJECT NUMBER A792	
			5d. TASK NUMBER 620	
			5e. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Advanced Training Methods Research Unit 5001 Eisenhower Avenue Alexandria, VA 22333-5600			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue  Attn: TAPC-ARI-II Alexandria, VA 22333-5600			10. MONITOR ACRONYM ARI	
			11. MONITOR REPORT NUMBER Technical Report 1133	
12. DISTRIBUTION/AVAILABILITY STATEMENT  Approved for public release; distribution is unlimited.				
13. SUPPLEMENTARY NOTES				
<p><b>14. ABSTRACT (Maximum 200 words):</b></p> <p>Students linked into a learning environment over the Internet may learn topics better by collaborating to create questions and answers. In research reported here, students used a learning aid for collaborative question generation called Army TEAMThink, a commercial program modified for Army use under a TRADOC Delivery Order contract. Research was done at three US Army schools to assess the quality or doctrinal correctness of questions and answers generated by students and to measure any learning benefit. Students first completed a tutorial on how to write effective multiple-choice questions. Next, students wrote questions and reviewed questions written by other students. Based on the feedback from the reviews, authors were allowed to modify their own questions. Finally, students took a test of the questions that had been developed by students using the learning aid. Army subject matter experts judged that most of the questions developed were considered acceptable and could be repurposed for use in course exams. A majority of the question feedback was constructive, indicating that the collaborative process was helpful. Students who went through the process scored higher on a test of novel questions than those who did not use Army TEAMThink. They also scored higher than students who went through the process on a different topic from the test topic, demonstrating a moderate learning effect. The general finding of this research about a collaborative question-generation approach is that instructors can accumulate quality multiple-choice questions and monitor student comprehension, and students have an additional opportunity for better learning.</p>				
<p><b>15. SUBJECT TERMS</b> training, educational technology, distributed learning, distance learning, collaboration, online instruction, web-based instruction, question generation, questions</p>				
SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	31
			21. RESPONSIBLE PERSON (Name and Telephone Number) James Belanich 703/617-2362	

**This Page Intentionally  
Left Blank**

**Technical Report 1133**

**Web-Based Collaborative Learning: An Assessment  
Of a Question-Generation Approach**

**James Belanich  
Robert A. Wisher  
U.S. Army Research Institute**

**Kara L. Orvis  
George Mason University  
Consortium Research Fellows Program**

**Advanced Training Methods Research Unit  
Franklin L. Moses, Chief**

**U.S. Army Research Institute for the Behavioral and Social Sciences  
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600**

**January 2003**

---

**Army Project Number  
20633007A792**

**Personnel, Performance  
and Training**

Approved for public release; distribution is unlimited.

**This Page Intentionally  
Left Blank**

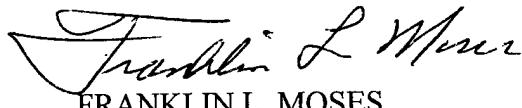
## FOREWORD

---

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is examining the use of distance-learning technologies for use by soldiers. The WEBTRAIN project seeks to provide guidance to the Army as it transforms from a classroom-centric method of instruction to one that is more soldier-centric and collaborative.

With the U.S. Army transforming many of its courses to a distributed-learning format, new methodologies are needed to take full advantage of current technology. A Memorandum of Agreement between TRADOC and ARI to evaluate commercial-off-the-shelf products led to this research to assess the potential of ARMY TEAMThink, a web-based program modified from a commercial product for U.S. Army use. The modifications were increased instructor control and the inclusion of a tutorial on question writing.

The findings of this report are relevant to course designers who may incorporate collaborative question generation as a learning tool. The results and recommendations of this research were presented to TRADOC on 17 October 2002, and also to representatives from the Total Force Advanced Distributed-learning Action Team on 11 April 2002, at the APA Midyear Symposium of Divisions 19 on 21 March 8, 2002, and at the E-Learning Conference and Expo on 10 April 2002.



FRANKLIN L. MOSES  
Acting Technical Director

## ACKNOWLEDGEMENTS

The authors would like to thank Tatana Olson and Jared Abramson for their efforts related to data collection and data analysis. Thanks are also given to Susan Thomas of the OMEM School at Redstone Arsenal, Helen Rimely of Fort Huachuca, and CPT Storm Reynolds and MAJ Steve Cross of Fort Leonard Wood for their help in coordinating data collection. James Dees of TRADOC, J. Douglas Dressel of ARI, and Art Graesser of the Memphis University deserve thanks for their assistance in reviewing earlier versions of the report. In addition, we would like to thank Jim Mazzu and Charles Dornbush of Athenium LLC for their technical assistance with the Army TEAMThink program.

## **EXECUTIVE SUMMARY**

---

### **Research Requirement:**

The Army is transforming many of its courses to a distributed-learning format. One criticism of distributed-learning is that students sometimes feel isolated from other learners. A method to remedy this is for students to work collaboratively, which leads to the exchange of knowledge through learner-learner interactions. The development of questions has been shown in previous research to be an effective learning tool. The beneficial components of learner-learner collaboration and question generation were combined in a software concept called Army TEAMThink<sup>1</sup>, a web-based asynchronous program adopted from a commercial product for US Army use. In the current research, the effects of the Army TEAMThink concept were assessed.

### **Method:**

The Army TEAMThink program was used at three TRADOC schools: the Engineer School; the Ordnance Munitions and Electronics Maintenance School; and the Intelligence School. Students were first required to complete a tutorial that instructed them on how to write effective questions. Then, the students, divided into separate teams, used a question template to write a multiple-choice question. Next, students answered questions that other students on their team had generated and provided feedback to the authors. Team members were then given an opportunity to modify their questions. Finally, students took part in a TEAM Challenge by answering questions developed by students from another team; this was the first time that the students had been exposed to these specific questions, although the questions were on the same course topic.

### **Findings:**

Subject matter experts rated approximately 75 percent of the questions developed as doctrinally correct, indicating that most of the questions could be used in a quiz or test. A comparison of the types of questions written by students who went through the tutorial and students who did not, revealed no difference in the semantic, conceptual and pragmatic aspects of questions developed. This indicates that the tutorial did not lead students to write questions that required a deeper level of understanding. A majority of the comments made to the questions were constructive, indicating that the collaborative process promoted by Army TEAMThink was helpful.

Students from the Engineer School who went through the Army TEAMThink process scored significantly higher ( $p < .01$ ) on a test consisting of questions generated during the Army TEAMThink exercise than students who did not go through the process (a mean increase of 7.7 percent), demonstrating a learning effect (effect size = 0.73). For students from the Ordnance Munitions and Electronics Maintenance School who took two different TEAM Challenges, one based on the topic that they used during the Army TEAMThink exercise and another based on a

---

<sup>1</sup> The commercial product TEAMThink was modified for use by TRADOC and for the purpose of supporting current research about the viability of such a concept for Army use. The use of Army TEAMThink does not constitute endorsement of the product by the US Army or the US Department of Defense.

separate topic from their course, a comparison of the two scores showed a mean increase of 7.3 percent (not statistically significant,  $p=0.16$ , effect size = 0.30).

#### Utilization of Findings:

The results of this research provide information about the utility of collaborative question generation software as a learning aid. The findings should particularly be useful to course designers who may develop training exercises that incorporate collaborative question generation as a learning tool.

# **WEB-BASED COLLABORATIVE LEARNING: AN ASSESSMENT OF A QUESTION-GENERATION APPROACH**

<b>CONTENTS</b>	<b>Page</b>
<b>INTRODUCTION.....</b>	<b>1</b>
ARMY DISTANCE-LEARNING.....	1
QUESTION GENERATION AS A FORM OF LEARNING.....	2
ONLINE COLLABORATION.....	3
<b>METHOD .....</b>	<b>5</b>
PARTICIPANTS.....	5
APPARATUS .....	6
PROCEDURE .....	6
<b>RESULTS .....</b>	<b>8</b>
QUALITY OF QUESTIONS .....	8
LEVEL OF QUESTIONS .....	9
PERCENT OF QUESTIONS ANSWERED CORRECTLY .....	10
COMMENTS TO QUESTIONS .....	12
SURVEY QUESTIONS - REACTIONS TO ARMY TEAMTHINK EXPERIENCE .....	13
<b>DISCUSSION .....</b>	<b>13</b>
QUESTION DEVELOPMENT .....	13
LEVEL OF QUESTIONS .....	14
CONSTRUCTIVE COMMENTS.....	14
LEARNING PROCESS .....	15
PARTICIPANT REACTIONS .....	16
LIMITATIONS.....	16
<b>CONCLUSION .....</b>	<b>17</b>
<b>REFERENCES.....</b>	<b>18</b>
<b>APPENDIX A.....</b>	<b>21</b>

## **Web-Based Collaborative Learning: An Assessment of a Question-Generation Approach**

### **Introduction**

As the US Army transforms many of its courses to a distributed-learning (DL) format, new technologies need to be developed to take full advantage of this learning environment. The types of training and education considered for distributed delivery cover a wide range of topics, including technical skills, collaborative tasks, skills associated with operating and maintaining equipment, and operating practices. Because the Army's goal is to convert over 500 courses to a DL format by 2010 (TRADOC, 1999), new training tools will have a far-reaching impact on Army training. Traditional instruction methods alone will not be adequate to effectively train today's soldiers in a broad range of skills; novel approaches to training are needed (Abell, 2000). Recent research has shown that the development of questions promotes increased understanding and comprehension in specific subject matter, and may be effectively integrated into a distance-learning environment (Graesser & Wisher, 2001). This report looks at the efficacy of combining on-line question generation and collaboration to improve learning.

#### **Army Distance-learning**

The use of DL by the Army and other services is presently in a growth phase. The Army Distance-learning Plan (TADLP) and the National Guard's Distributed Training Technology Program (DTTP) intend to provide easy access to Web-based courses for soldiers. The current plans call for the development of digital training facilities (DTF) located throughout the United States and the world with the goal of having a DTF within 50 miles of 95 percent of soldiers in CONUS (Program Management Office, The Army Distance-learning Program, 2000). Easy access to web-based courses has the potential to enhance Army organizational performance dramatically by increasing access to training. The development of effective training tools will only further exploit the efficiency of DL by the Army.

In general, distributed-learning has been shown to be at least as effective as traditional classroom instruction (Metzke, Redding, & Fletcher, 1996; Howard, 1992; Navarro & Shoemaker, 2000). Based on a review of 30 studies comparing the time needed to complete courses, Metzke, Redding, & Fletcher (1996) found that students needed an average of 30 percent less time to complete computer-assisted and computer-managed instruction compared to traditional instruction for achieving the same level of competency. Metzke, et al. (1996) also found across various instructional settings, that computer-based instruction resulted in a higher level of achievement compared to traditional instruction, with an average increase of .5 standard deviations. However, another study on Web-based instruction showed a more limited effect of approximately .3 standard deviations (Wisher & Olson, in review). Improved tools for DL and computer-based instruction may continue this positive trend of increased efficiency in instruction, showing both faster completion rates and higher achievement than is presently available.

With the Army providing over 100 million hours of training to active duty and reserve personnel annually, the potential savings through the implementation of DL is substantial (Defense Manpower Data Center, 2002). Distributed-learning and computer-based instruction

have the possibility of increasing personnel qualifications in the unit and reducing the impact of skill decay by making training available when and where it is required. Distributed-learning is widely recognized as the method of choice for reducing costs while increasing flexibility, course access and the number of learners reached. The flexibility of distributed-learning is also attractive to the Army for mission execution; for example, the ability to conduct pre-deployment, mission specific training under the tutelage of dispersed subject matter experts can result in faster preparation for contingencies. In addition, the use of DL can level the playing field for the Reserve Component and for geographically remote organizations and learners, by providing a standardized learning experience without walls or barriers (Freeman, Wisher, Curnow, & Morris, 1999). All of these factors make DL an attractive option for the Army.

Although DL has already been shown to save time and money while increasing effectiveness, DL still is in a formative stage. With continued study of the features of DL that are effective, even more advantages may be realized (Abell, 2000). Two areas of promising research are the use of question generation (Graesser & Wisher, 2001) and the use of collaborative tools (Bonk & Wisher, 2000) in an online environment.

### Question Generation as a Form of Learning

The study of question generation and question-based learning has primarily focused on the asking of sincere information seeking (SIS) questions (Graesser, Person, & Huber, 1992; Graesser & Wisher, 2001; Minstrell, 1999). With SIS questions, the questioner does not know the answer but expects that the person being asked will provide an answer. In such a scenario, the most obvious result of asking questions is the receiving of answers. When a question is sufficiently answered, the questioner acquires knowledge.

The process of learning, however, does not end with the answer to a question. When a questioner receives a useful answer, he or she is more likely to ask more questions, which in turn leads to the development of an inquiring learning environment (Minstrell, 1999). The asking and answering of questions, therefore, builds upon itself and produces common ground between participants, where they effectively share information (Graesser, Person, & Huber, 1992).

To benefit from questions, a person does not need to be either the questioner or the answerer, but merely have access to observe the question and the answer. A study by Shavelson, Berliner, Ravitch, and Loedding (1974) found that being exposed to questions on a topic increased participant's subsequent accuracy while answering entirely different questions on the same topic area. This occurred even when answers were not given.

The process of thinking about a topic and producing a SIS question requires active processing of information; therefore, the asking of questions should lead to more in-depth learning (Graesser & Wisher, 2001). To develop a question, several metacognitive tasks must be completed. First, the questioner must identify what information they do not know. Based on the lack of information, they must develop a question in a way that would be understandable to others. This development requires processing at a more in-depth level than merely sequencing words at a surface level. It requires the latent semantic understanding of terms that are commonly used by the intended answerer. It also requires the questioner to gauge the level of

understanding of the person who is asked the question. This increased meta-cognitive load may be why some people have difficulty generating effective questions, even when they understand the content area (Miyake & Norman, 1979; Otero & Graesser, 2001). All of these steps require deeper processing of information, which in turn leads to greater comprehension (Craik & Lockhart, 1972).

In addition to seeking information from others, asking questions provides information about the questioner. The questions asked by a person about a particular topic allow others to assess the level of understanding of the topic by the questioner (Swartz, 1987). For example, a person must have at least some knowledge about a topic in order to ask an in-depth question. In addition, the questions identify content areas where the questioner has incomplete information. Used in this manner, the asking of questions can provide information about the questioner, identifying what the questioner knows and what the questioner does not fully comprehend.

Not all questions, however, accomplish the intended goals of the questioner. Sometimes the answerer does not fully understand the intent of the question, and therefore provides an inappropriate response. Graesser and Wisher (2001) have shown that learning about the functional components of questions and how to ask an effective question improves question-asking skills and subsequent learning. When a question clearly expresses the information being sought, the answer is more likely to be helpful and instructive.

Most of the research regarding question asking has focused on the asking of questions in a face-to-face setting (Graesser & Wisher, 2001). There have been some studies, however, that assessed question asking in a web-based setting. One such study (Blanchette, 2001) found that questions are developed differently in a face-to-face setting than in an online learning environment. In an online environment, the questions tend to be at higher cognitive levels than face-to-face, which is probably due to the ability to edit and examine the questions while typing.

### Online Collaboration

The American Psychological Association acknowledged the importance of social interactions in the learning environment in their “Learner-Centered Psychological Principles” (American Psychological Association, 1997). These principles reiterated the importance of learners collaborating to construct knowledge. Some additional principles of learner-centered instruction, such as directed goals, metacognitive skills, and motivating activities are supplementary to collaboration. The Army TEAMThink program, to be discussed in detail below, attempted to apply these learning principles via on-line collaboration and question generation.

Through collaborative tasks, such as discussing, summarizing, clarifying, and integrating course content into an overall framework, learners acquire knowledge and gain deeper understanding (Deatz & Campbell, 2001). Collaborating with fellow learners allows for the synergistic building of knowledge; the strength of one team member can be shared with the other team members who do not have the same strengths. This form of collaboration, which applies social constructivist principles, leads to a productive learning environment (Palincsar, 1998). Numerous studies have shown that the use of collaborative learning exercises leads to improved

knowledge acquisition when compared to traditional classroom instruction without collaborative exercises (Alkhateeb & Jumaa, 2002; Bonk & Wisher, 2000; Fisher & Coleman, 2001-2002; Moore, 2001).

One form of instruction where the effective use of collaborative tools is emerging is web-based instruction (Duffy, Dueber, & Hawley, 1998). People in distance-learning courses often feel isolated because of a lack of natural interaction with class members (Muilenburg & Berge, 2001). In a DL environment where learners don't have the same opportunity to interact face-to-face with other students as in a traditional classroom, the process of learning can become lonely and noninteractive. For example, students are not able to bounce ideas off of one another after class, or ask the student sitting next to them for clarification during a lecture. Navarro & Shoemaker (2000) found that many learners in DL courses think that more learner-learner interaction should be designed into online courses to combat the problems associated with isolated learners in DL courses. Collaborative online tools, therefore, are needed to increase learner-learner interaction that motivates students in a DL environment (Fisher & Coleman, 2001-2002).

Online collaboration can take place synchronously, when all students are present at the same time, or asynchronously, when the participants are online at different times. With synchronous collaboration, students feel connected with each other because of the immediacy of the communication (Kang, 1998). Asynchronous collaboration allows for more flexibility on the participant's time of access, and may be preferable in some situations because it allows students to interact even when they cannot be on-line at the same time. Both synchronous and asynchronous collaboration tools are important features of a DL course (Duffy, et al., 1998).

According to Garrison, Anderson, and Archer (2001), another important aspect of online collaboration in a DL environment is that it usually is structured to promote inquiry. This includes defining problems or topics, trying multiple solutions, feedback, and reflecting on the process and outcome. Garrison, et al. found that with asynchronous text-based conferencing, most of the communication could be identified as exploration, the seeking of new information.

According to Moore (1989), there are three types of interactions in instructional settings: learner-instructor, learner-learner, and learner-content. All three types of interaction promote learning, but in traditional learning environments, the emphasis is usually on learner-instructor and learner-content interactions because it is generally assumed that instruction flows towards the learner. Social constructivist theories of learning, however, place a high level of importance on learner-learner interaction. The importance of learner-learner interaction through the use of text messaging was demonstrated by Kang (1998), where 85 percent of text messaging during a graduate-level course was learner-learner, showing a high level of learner-learner interaction. The students felt that the use of text messaging allowed relationships to develop and to improve social bonds among students. In addition, an analysis of the students' text suggests that they think more when writing text messages compared to verbal interactions. In text messaging, students have time to ponder, edit, and express themselves, which leads to more thoughtful and content-rich communication. Savery's (1998) findings were similar to those of Kang, showing that on-line collaborative writing fosters learning. Savery found that collaborative writing promotes cognitive and metacognitive skills through the process of constructing knowledge.

One form of collaborative writing exercise that might be beneficial to learners is to develop questions on a specific topic and test one another with those questions. In the current study, the process of collaboration and question-based learning in a web-based setting was analyzed. The software used, Army TEAMThink™, was a web-based platform where a team of users collaborated to develop questions.<sup>2</sup> Army TEAMThink, produced by Athenium LLC, is a modification of a commercial product for use by the Army under a contract with TRADOC. The effects of collaborative question writing on user outcomes such as test performance, quality of questions developed, interaction, and satisfaction were assessed.

## Method

### Participants

The participants were students at three Army schools: the Engineer School at Fort Leonard Wood, Missouri; the Ordnance Munitions and Electronic Maintenance School at Redstone Arsenal, Alabama; and the Intelligence School at Fort Huachuca, Arizona. A total of 199 participants went through the Army TEAMThink exercises, as shown in Table 1. To complete the Army TEAMThink exercises, the participants needed a total of about 3-4 hours over two sessions.

At the Engineer School, participants were students in the Engineer Captains Career Course. Within the Engineer School, 104 soldiers were assigned to the experimental group that participated in the Army TEAMThink exercise and 108 soldiers were assigned to the control group that did not participate in the Army TEAMThink exercise. At the Ordnance Munitions and Electronic Maintenance School, participants were students in either the Basic Non-commissioned Officers Course (BNCOC) or the Advanced Non-commissioned Officers Course (ANCOC) for training as an ammunition specialist (MOS 55B). Within the Ordnance Munitions and Electronic Maintenance School, a total of 40 soldiers, divided into five groups, participated in Army TEAMThink exercises. The five smaller groups were needed to accommodate these soldiers taking two separate tests, one with questions developed within their group and topic and another with questions developed by soldiers in other groups who used a different topic. At the Intelligence School, the 55 soldiers who participated were students in the Military Intelligence Captains Career Course, and they all used Army TEAMThink.

Table 1. The number of participants at each location.

Engineer School		Ordnance Munitions and Electronic Maintenance School		Intelligence School	
Exp. Group	N= 104	Group 1	N = 8	Total	N = 55
Cont. Group	N = 108	Group 2	N = 9		
		Group 3	N = 6		
		Group 4	N = 11		
		Group 5	N = 6		

<sup>2</sup> The use of trademark or copyright material does not imply endorsement of a product by the US Army Research Institute for the Behavioral and Social Sciences or by the Department of the Army or the Department of Defense..

## Apparatus

The participants used computers with Internet connections to access Army TEAMThink. For these experimental sessions, each of the participating schools provided access to the computers. The Army TEAMThink program and all data were stored on a server at Athenium in Boston, MA. In the current experiment, Army TEAMThink was used in a traditional classroom setting due to experimental and scheduling constraints. Students in this experiment participated in only one iteration of Army TEAMThink; therefore, we are unable to generalize our findings to multiple exposures to the process.

## Procedure

*Tutorial.* All soldiers first completed an Internet-based, self-paced tutorial, which was about 40 minutes in length, instructing them on how to author effective multiple-choice questions. The tutorial focused on developing different parts of a question. The topics of the tutorial covered the basic components of multiple-choice questions including context, stem, key, and distractors. The context is supplemental information in the question that provides situational background for the question. The stem is the actual question. The key is the correct answer option. The distractors are the incorrect answer options. The principles expressed in the tutorial were based on research done by the U. S. Army Research Institute for the Behavioral and Social Sciences and the Army Training and Doctrine Command. For additional information regarding the use of question generation as a training tool, see Graesser and Wisher (2001).

The tutorial was divided into three modules that focused on the following areas: a) learning through questions; b) writing properly formatted question stems; and c) creating a set of answers that includes only one correct answer and three appropriate distractors. During each module, appropriate examples of question stems or answer sets were presented as models. Each module was followed by a brief quiz on the information provided in the module. The requirement was to cycle through a module until achieving a 100 percent correct score before continuing to the next module. After all three modules were completed, a final assessment covering all modules was presented.

*Question Collaboration.* Immediately after completing the tutorial, students were divided into teams and asked to individually compose a question. The size of the teams varied from 3 to 24 members. The Army TEAMThink program provided a template into which students wrote multiple-choice questions along with corresponding correct answers and distractors. The template included an area to write the question stem and separate areas to write the correct answer and distractors, as shown in Figure 1. Additional sections of the template included what the correct answer was, a space for the rationale for the question, and a section where the question writer could include any reference material that might be helpful in answering the question.

All participants were given particular topics on which to write their questions based on the course they were taking. In all locations, half of the participants in a specific class wrote

questions on one topic (e.g., military decision making) while the other half wrote questions on a different topic (e.g., storing of munitions). Because there were a small number of students in each group from the Ordnance Munitions and Electronic Maintenance School, the soldiers were instructed to write three questions, so that there would be a sufficient number of questions for a test. In the other locations with larger groups the soldiers were asked to write only one question. As a general guideline, students were asked to write questions in their respective topics that they thought approximately 70 percent of the class could answer correctly.

After writing their own questions, team members reviewed the questions developed by the other members of their team. The participants read and answered each question their teammates wrote. After answering each question, students were shown the answer identified as being correct by the question's author. Participants then were prompted to write comments, providing feedback to the author of the question.

**Review Team Questions - Microsoft Internet Explorer provided by MSN**

File Edit View Favorites Tools Help Address:

Question 2 of 7 \* not reviewed

Last updated by user E129 06:45PM, 11-Apr-00 (ET)

**Question**  
You have a radio system built to communicate with your neighbors over in East Berlin. It uses 50 kW to transmit at 1800 MHz. The receiver is only 100 m away. The Germans, however, wish to build their wall directly between you and your neighbors, only 40 m distant from the receiver. If the complex receiver on the East German side can detect signals at as low as -20 dBm, how high does the wall need to be to terminate communication? Assume unity gain in both tx and rx.

**Possible Answers**  
the selected answer is the one that the author chose

a) <input type="radio"/> C	1.74 m
b) <input type="radio"/> C	2.84 m
c) <input checked="" type="radio"/> C	2.89 m
d) <input type="radio"/> C	4.002 m
e) <input type="radio"/> C	Too far, can't communicate anyway.

**Attachment**  
(none)

**Rationale**  
First, find the free space power received from eqn 3.15. This yields .88 mW or -0.56 dBm.  
Since the receiver can receive signals at -20 dBm, this allows for -19.44 dB in loss. Looking at the graph on p. 97, we see that this amount of loss will yield a Fresnel Diffraction parameter between 1 and 2.4. Using equation 3.61.d, we find this parameter  $v = 2.001$ .  
Then using  $d_1 = 40$ ,  $d_2 = 60$ ,  $\lambda = 1/6$ , we can use eqn. 3.56 to find  $h$ . It gives 2.831 m. This fulfills the requirements that  $h \ll d_1$ ,  $d_2$  and  $h \gg \lambda$ . The wall must then be at least this value to obstruct communication. The answer is therefore (b).

**References**  
You have not commented on this question.

**SUBMIT** Question ID: 62

**Comments for Team Question 2** Jump to most recent comments

**Author/Review Question phase (previous comments)**

User ID: A139 Time/Date: 06:43PM, 11-Apr-00 (ET)  
 Agree  Disagree  
 Using the exact value of 2.83m actually produces a received power above the receivable level (-49.9963dB). Because the exact number that I calculated was a little bit above 2.83m (2.831m), I picked 2.89m. I know it's picky, but it's true...

User ID: A139 Time/Date: 06:20PM, 11-Apr-00 (ET)  
 Agree  Disagree  
 This is a good question but needs some tweaking. User A139's comment is valid. In addition, please use dB and dBm properly. In the question and rationale and user A139's comment, dB's are confused with dBm. Please revise.

Figure 1. The template for writing questions, answer stems, and comments.

When the members of a team completed the review and commenting process, they were directed to return to their own questions. At this point, participants were able to read the comments written about their question and make revisions if they desired. The reviewers of a particular question could also read comments posted by other team members. Once the authors finalized the questions, the course instructors assessed the questions for doctrinal compliance. Instructors accepted questions that they thought addressed the learning objectives of the course, adhered to Army doctrine, and were formatted properly. For examples of accepted questions and comments to questions from all three locations, see Appendix A.

*TEAMChallenge*. The final step of the Army TEAMThink process was for the participants to take an Internet-based test (TEAMChallenge) of the questions developed by another team that wrote questions on the same topic area. The questions that appeared on the TEAMChallenge were only questions that were deemed acceptable (adhered to Army doctrine) by the course instructors. Participants taking a TEAMChallenge had not previously seen the questions they were given, since a separate team developed them. Although they had participated in Army TEAMThink on the same topic, no team answered questions during the TEAMChallenge that they had developed and reviewed earlier. They answered questions on the same topic, but only ones developed by another team. The number of items on the test varied based on the number of questions developed and the number of questions accepted by the course instructor.

The participants from the Ordnance Munitions and Electronic Maintenance School took two TEAMChallenges. The first TEAMChallenge consisted of questions from the topic that they used during their Army TEAMThink exercise, and the second TEAMChallenge consisted of questions from another course topic on which they did not go through the Army TEAMThink process. The instructor, however, had recently covered both topics. This allowed for the comparison of how well participants answered questions in the topic area they used for Army TEAMThink and in a topic that they did not use for Army TEAMThink.

## **RESULTS**

### Quality of Questions

Across all three locations, 336 questions were written. Subject matter experts (course instructors) judged 77 percent of the questions as doctrinally correct and acceptable, as shown in Table 2,. Across locations, the percent of questions deemed doctrinally correct was similar. For the participants from the Engineer School, 81 percent of the questions were accepted. For both the Ordnance Munitions and Electronic Maintenance and the Intelligence School, 75 percent of the questions were accepted.

Table 2. The number of questions authored, questions accepted, and the percent of acceptable questions across all locations.

	Authored	Accepted	Percent
Engineer School	101	82	81%
OM&EM School	180	135	75%
Intelligence School	55	41	75%
Total	336	258	77%

### Level of Questions

According to Graesser, Person, and Huber (1992) and Graesser and Person (1994), there are 18 different types of questions based on semantic, conceptual, and pragmatic aspects. Of the 18 types of questions possible, Graesser et al. (1992) organized the question types into three levels: shallow, in-depth, and other. Five types were considered shallow, eleven types were considered in-depth, and the remaining two types fall into an “other” level. For this research, only the shallow and in-depth levels were considered. The five question types in the shallow level were: *verification*, yes/no answers; *disjunctive*, choices are given to select from; *concept completion*, an incomplete concept is stated with the answer being the missing component of the concept; *feature specification*, requires knowing the components of the item/concept; and quantification, the answer must be a numerical value. The eleven question types in the in-depth level were: *definition*, meaning of a term, as in a dictionary entry; *comparison*, items are compared or contrasted; *example*, a representative member of a category/concept is required; *interpretational*, an inference of available data is required; *judgmental*, requires the formation of an opinion; *antecedent*, answer is the set of events that led to a known outcome; *consequence*, answer is an outcome caused by stated precursors; *goal orientation*, the reason (motivation) for acting a particular way in pursuit of a goal; *enablement*, the object or resource that allows an agent to perform an action; *instrumental/procedural*, steps that lead to the achievement of a goal; and *expectational*, reason why an expected event did not occur.

Using the Graesser, Person, and Huber (1992) coding scheme, all of the questions were categorized by a computer program custom designed to identify question types. The software used was an experimental question classifier developed for the AutoTutor Project at the University of Memphis (A. Olney, personal communication, January 26, 2002).

The types of questions were compared across students from the Engineers course. Two sets of questions were assessed from the Engineering course, one was the experimental group that went through the Army TEAMThink process and the other set of questions was from students who did not participate in the Army TEAMThink exercise, but just wrote questions. This second set of Engineers was used as a control by writing questions without completing the tutorial and without participating in the Army TEAMThink exercise.

The relative rate for the types of questions was similar for the experimental and control groups; a correlation between groups was statistically significant (Spearman,  $r=0.85$ ,  $p<0.01$ ). A summary of the question types authored, using the shallow and deep categories, is presented in Table 3. For the experimental group, 70 percent of the questions were shallow and 30 percent were deep. For the control group, 67 percent of the questions were shallow and 33 percent were

deep. This indicates that the types of questions did not differ for soldiers who were exposed to the tutorial and wrote their questions using Army TEAMThink versus soldiers who were not exposed to the tutorial and wrote their questions with pen and paper.

Table 3. The percent of questions written by the experimental and control groups of students from the Engineering School that were shallow or deep.

	Experimental Group	Control Group
Shallow	70%	67%
Deep	30%	33%

#### Percent of Questions Answered Correctly

To determine if a learning effect occurred due to the use of the Army TEAMThink program, two different analyses were conducted. For both analyses, the percent of questions answered correctly was the dependent variable. In the first analysis, the percent of questions answered correctly by the soldiers in the control and experimental groups from the Engineer school were compared. In the second analysis, two sets of questions were used to assess the soldiers from the OM&EM School; these included questions about the topic that they used for the Army TEAMThink experiment and a topic that they did not use for Army TEAMThink, but covered in class. The percent of questions answered correctly was used as a measure since the number of questions in each test varied.

*Engineer School.* For the Engineer School, the percent of questions answered correctly during the TEAMChallenge was compared between the experimental and control groups. Although these groups answered different questions, the comparison of scores is valid because the depth of questions across the two groups was similar (see analysis above using the Graesser, Person, & Huber coding scheme). For the experimental group, the participants answered 76.1 percent of the questions correctly, as shown in Figure 2. For the control group, the participants answered only 68.4 percent of the questions correctly. The 7.7 percent difference between means was statistically significant (independent samples  $t=3.76$ ,  $df=210$ ,  $p<.01$ ). The effects size for this difference was 0.73.

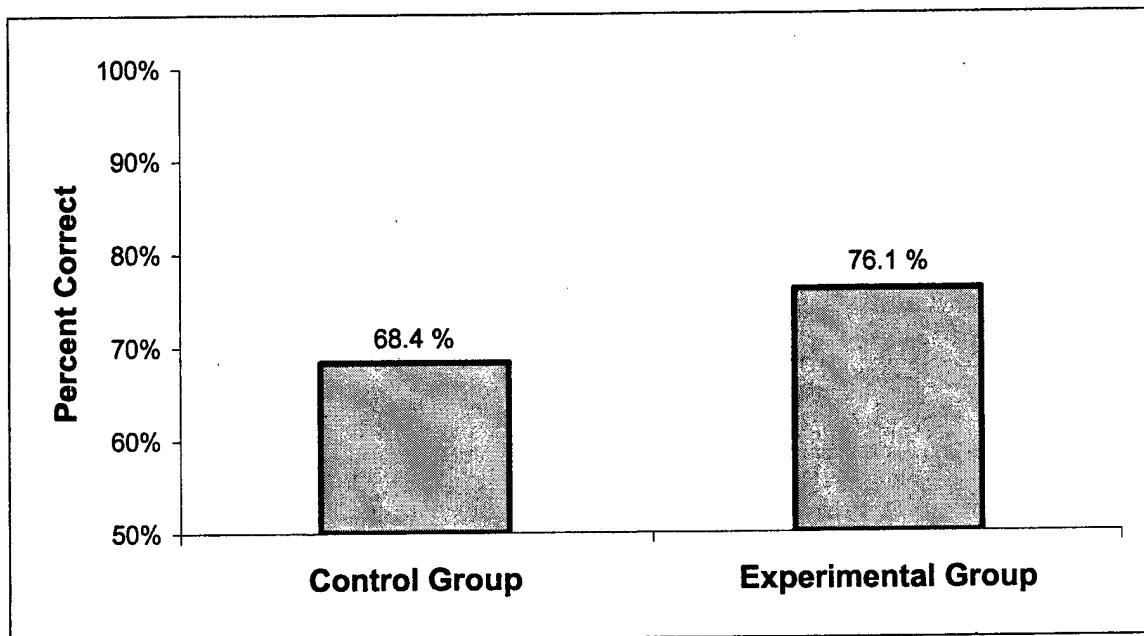


Figure 2. The percent of questions answered correctly for the control and experimental groups from the Engineer School.

*OM&EM School.* The students from the OM&EM School took two TEAMChallenges, the first was with questions in the same topic area as they developed their own questions (“within” on Table 4) and the second in a different topic area (“across” on Table 4). Both of the topics were covered during the actual course that the students were currently taking, but any particular soldier used only one of the topics during the development of questions. In both of the tests, it was the first time participants were exposed to the questions. Participants averaged 7.3 percent higher scores on TEAMChallenges on topics in which they had Army TEAMThink experience versus topics on which they had no Army TEAMThink experience. This suggests that the use of Army TEAMThink in a particular topic area led to an increased ability to answer questions in the same topic area, but these results were not statistically significant ( $p=0.16$ ). The effects size for this difference was 0.30.

Table 4. A comparison of scores of the participants from the five OM&EM School groups that participated in two TEAMChallenges (\* indicates  $p<.05$  level).

	Group 1	Group 2	Group 3	Group 4	Group 5	ALL
Avg. within	77.8%	71.8%	91.7%	77.7%	55.7%	75.2%
Avg. across	55.9%	70.9%	65.7%	73.4%	71.7%	67.9%
Difference	21.9%	0.9%	26.0%	4.3%	-16.0%	7.3%
N=	8	9	6	11	6	40
t =	2.46	0.06	2.70	0.51	-1.58	1.42
p =	0.043 *	0.952	0.043 *	0.622	0.176	0.163

One reason for the lack of a statistically significant difference was that after reading the comments from the “within” test, the question authors had the opportunity to edit questions before the presentation of the “across” test was administered. For example, a poorly written, confusing question that appeared in the “within” test may have been edited to make it clearer before it was presented in the “across” test. For the questions that were not edited, the likelihood of getting a particular question correct decreased 1.8% from the within topic test to the across topic test. For the questions that were edited, the likelihood of getting a particular question correct increased 2.7% from the within topic test to the across topic test. These scores compare the likelihood of a correct response when questions were presented to different groups: a) participants who were part of an Army TEAMThink exercise in the same topic area, and b) participants who were part of an Army TEAMThink exercise in a different topic area. Some participants did not complete the TEAMChallenge and some took only one TEAMChallenge. The questions that were edited might have been clarified, and, therefore, easier to answer. This may have muted the effect of “within” scores being higher than “across” scores, and may account for the lack of a significant difference between the “within” and “across” scores.

### Comments to Questions

The comments made to the questions were assessed using a 5-point classification system. The types of comments were: constructive, correct/incorrect, value statement, unrelated, and nonsense/trivial. A constructive comment was one that mentioned question content, answer content and/or specifically discussed how the question can be improved. A correct/incorrect comment was a comment that stated whether question and answers are correct or incorrect without mentioning content. A value statement comment stated either a positive or negative position towards the question with no mention of content (e.g., “good question”). An unrelated comment did not relate to the question content, and did not state a clear indication of the commenter’s view of the question (e.g., “who wrote this question?”). Trivial/Nonsense comments were not understandable (e.g., “xasgf”). This comment-coding scheme was developed for this research. Using a random sample of 52 comments, an inter-rater agreement score [agreements/observations] of 0.92 was obtained (Cohen’s Kappa = 0.9).

As shown in Table 5, a majority of comments (57%) were constructive, and an additional 6 and 17 percent fell under the correct/incorrect and value statement categories, respectively. This indicates that 80 percent of comments provided at least some level of useful feedback regarding the question. The remaining 20 percent of the comments were either unrelated (18%) or trivial/nonsense (2%). As mentioned earlier, participants were able to write comments during both the collaboration phase and during the TEAMChallenge. There were no differences in both the frequency and the types of comments written during the two phases or between schools.

Table 5. The actual number and percentage of each type of comment written during the question authoring phase, TEAMChallenge phase, and total for both phases.

Type of comments	Author Phase	%	Challenge Phase	%	Total	%
Constructive	228	56	202	57	430	57
Correct/Incorrect	20	5	25	7	45	6
Value Statement	79	19	53	15	132	17
Unrelated	76	19	63	18	139	18
Trivial/Nonsense	4	1	9	3	13	2
Total	407	54	352	46	759	100

#### Survey Questions - reactions to Army TEAMThink experience

After the OM&EM School participants completed the TEAMChallenge, they were given a set of questions to gauge their perception of Army TEAMThink. The results of the participant reaction questions are shown in Table 6. The ratings regarding the tutorial were rather positive, while less than 50 percent of the participants reported that the interaction through comments was helpful.

Table 6. Participant reactions to Army TEAMThink.

87%	Rated the responsiveness of the Internet tutorial as “adequate” to “very good”
77%	Rated the Army TEAMThink website as “easy” or “very easy” to use
76%	Rated the tutorial as “adequate” to “very good”
45%	Stated that the comments their teammates made helped to develop better questions
33%	Stated that using Army TEAMThink helped them understand the topic better
29%	Stated that using Army TEAMThink was interesting

## DISCUSSION

#### Question Development

During this experiment, course instructors judged more than 250 of the 336 multiple-choice questions authored as acceptable. This indicates that the instructors could repurpose approximately 75 percent of the questions produced during this exercise, and use them for end-of-module tests. The generation of so many quality questions in this research implies that a question generation system could be used as a tool to generate a corpus of questions for use by training developers, consequently saving time and reducing course development costs. The development of questions for tests often takes a lot of preparation time for teachers and trainers (Moore, 2001). Furthermore, the software can calculate psychometric measures such as percent of individuals who answered each question correctly, providing information to rank questions based on difficulty. This would be helpful to a course designer considering computer adaptive testing.

The review of the questions and the determination that most of them were doctrinally correct was reassuring to the instructors. The instructors also reported the questions demonstrated students command of course content. This is congruent with Swartz (1987), who stated that a question might expose the level of understanding by the questioner. Since basic information on a topic is required to develop an effective question, perhaps instructors could use a question generation system to monitor student comprehension during a course. If most of the questions developed on a particular topic were acceptable, the instructor would know that the class in general has a clear understanding of the topic. If many of the questions were inaccurate or confusing, the instructor would know that a review of the topic with the class was necessary.

### Level of Questions

The questions developed by the participants in this experiment were assessed using the Graesser and Person (1994) coding scheme. This coding scheme categorizes questions based on the questioning mechanism they incorporate and on the type of knowledge the questions were designed to measure. The assessment indicated that the questions developed by the participants of the current research were at a level similar to questions asked during tutorial sessions for college students (Graesser & Person, 1994). The correlation of the types of questions was statistically significant (Spearman  $r=.74$ ,  $p<.001$ ) between the corpus of questions developed using the Army TEAMThink software and questions asked during a tutoring session with a college student (Graesser & Person, 1994). This suggests that the Army TEAMThink concept can be used as a study aid.

The questions created by the participants who went through the tutorial were similar, in terms of the Graesser and Person (1994) scale, to the questions created by the participants who did not go through the tutorial. Questions from both groups used the same questioning mechanisms with equal likelihood, regardless of the authors' exposure to the tutorial. This may be due to the years of exposure to multiple-choice questions prior to the tutorial. The 45-minute tutorial may not have been enough to overcome this past experience.

### Constructive Comments

A total of 759 comments were made in response to the 336 questions authored, for an average of 2.3 comments per question. In addition, a majority of the comments were constructive, providing useful feedback to the question's author, which implies positive interaction between team members. To write a constructive comment, participants needed to have some basic understanding of the question's topic and to process that information. The mention of question content in a majority of the comments indicates that the participants were thinking about the material. This process suggests active learning, which has been shown to improve comprehension (Craik & Lockhart, 1972; Deatz & Campbell, 2001; Graesser & Wisher, 2001). The writing of comments, therefore, may lead to greater understanding of the question's topic.

The need for learner-learner interaction in a DL environment is fundamental. The use of a collaborative program such as Army TEAMThink may therefore be an effective means to

promote learner-learner interaction, overcoming the isolated feeling that learners in distributed-learning courses have reported (Alkhateeb & Jumaa, 2002; Bonk & Wisher, 2000; Duffy, Dueber, & Hawley, 1998; Fisher & Coleman, 2001-2002; Kang, 1998; Moore, 1989).

### Learning Process

The social influence on learning (e.g., social interactions and communication with others) is being recognized as an important contributor to learning (Bonk & Kim, 1998). The collaboration that occurs during Army TEAMThink may aid in the learning process, as shown in the higher scores by the participants from the Engineer's School who went through Army TEAMThink versus the participants who just wrote questions and took a test. For the participants from the OM&EM School, when all of the participants' scores were pooled together, the difference between the "within" and "across" groups was not statistically significant, even though there was a difference of over 7 percent. The scores, however, for 2 of the 5 subgroups had means and standard deviations that distinguished them. These two subgroups did significantly better in the TEAMChallenge in the topic area that they used for their own questions than in the other topic area. This suggests that a limited learning effect may be evident after only one use of Army TEAMThink. The learning effects found, although not conclusive, do point in a positive direction. We cannot tell from this experimental data what the effects would be if Army TEAMThink was fully incorporated into a DL course and used multiple times during the course.

The underlying cause of the limited learning effect demonstrated in this research is not apparent. It may have come from the process of developing the questions, or of being exposed to questions related to the topic and debating them (Graesser & Wisher, 2001). According to Shavelson, Berliner, Ravitch, and Loedding (1974), exposure to questions on a topic may increase ability to correctly answer other questions in the same topic area. Another possible cause is that increasing social interaction leads to greater learning and involvement by students (Moore, 1989). Additionally, it might be the active process of writing comments about the questions (Deatz & Campbell, 2001; Graesser & Wisher, 2001). At this point, it would be pure speculation to identify the component, or components, of Army TEAMThink that appeared to promote learning.

The finding that the use of Army TEAMThink led to increased comprehension of the course content by students, while not always statistically significant, does have theoretical support. Deatz and Campbell (2001) proposed that learner control, performance feedback, and rehearsal of course content should all lead to increased comprehension. According to the "Learner-Centered Psychological Principles" proposed by the American Psychological Association, collaborative social-constructive interactions, such as those that occurred during the Army TEAMThink exercise, are an important component of the learning environment (American Psychological Association, 1997). The current research adds to a growing literature showing positive results for programs that promote collaborative learning (Alkhateeb & Jumaa, 2002; Bonk & Wisher, 2000; Fisher & Coleman, 2001-2002; Kang, 1998; Moore, 2001) and question-based learning (Graesser, Person, & Huber, 1992; Graesser & Wisher, 2001; Minstrell, 1999; Shavelson, Berliner, Ravitch, and Loedding, 1974; Swartz, 1987). The findings suggest that

collaborative question generation can be used as an asynchronous, web-based tool to increase learning by students.

### Participant Reactions

Some of the results from the participant surveys appear to contradict some of the results of the research. While the tutorial received high marks from the participants, there seemed to be little impact of the tutorial on the level of questions authored. While the results suggested that a limited learning effect might have occurred, less than 50 percent of the students felt that Army TEAMThink helped in understanding the topic and developing better questions. According to Alliger, Tannenbaum, Bennett, Traver and Shotland (1997), reaction measures to learning experiences do not necessarily correlate with learning outcomes. This may account for the disparity between participants' perceptions as to the value of their Army TEAMThink experience and their performance.

### Limitations

There are several issues that may have limited the effect of Army TEAMThink during this research. First, the participants in this research were all aware that their performance would not count toward their final grade, as explained in the instructions given. This may have led some of the participants to not seriously apply themselves during the experiment. If Army TEAMThink were incorporated into a course, different outcomes might unfold.

In the current research, the question development process was not repeated after the question authors received feedback on their initial questions. If a participant received some feedback about how to develop an effective question and also observed others writing effective questions, they may be more likely to write effective questions in the future. Repeated exposures might have an effect on the quality of question development over time.

The types of questions were based on Graesser and Person's (1994) scale, which focused on the question stem. While the Graesser and Person scale does have a high correlation with Bloom's taxonomy of educational objectives (Graesser & Person, 1994), this form of analysis does not consider factors like the use of distractors, question difficulty, or question clarity. Different results might have occurred if these other factors were assessed on an individual question level.

Finally, the participants were instructed to write questions that would probably result in about 70 percent of the participants answering correctly. Changing the level of difficulty of the questions may affect their quality. Surprisingly, participants were able to accurately estimate the overall percentage of students who answered their questions correctly. Across the three locations, the percentage of questions answered correctly by students was approximately 75 percent.

## **Conclusion**

The use of Army TEAMThink for collaborative question generation produced positive outcomes for both students and instructors. The benefits to the students were increased comprehension of the course material and a novel method of interacting with other students. The benefits to the instructor were the development of quality questions with minimal time investment and to monitor the topic comprehension by students.

The findings of the current research along with previous findings suggest that collaborative, question-generation exercises are effective learning tools. In the current research, students demonstrated an increase of approximately 7 percent in the questions they were able to correctly answer after using Army TEAMThink. The Army TEAMThink concept is based on the principles of collaborative and question-based learning. The attempt was to combine collaboration and question development to aid in the learning of course content. The positive results of this research might be specific to the Army TEAMThink program, or may be generalizable to collaboration and/or question-based learning in general.

The Army TEAMThink program was designed as a web-based program so that it could be incorporated into either a distributed-learning or a traditional course. In either design, an instructor would receive, as a by-product, a large number of questions, most of which could be easily reused during the course, or during subsequent iterations of the course. The development of these questions would help instructors save time in creating exams, and time is a highly regarded resource of DL instructors (Moore, 2001). An additional benefit to instructors is that through the implementation of Army TEAMThink, an instructor can gauge the level of comprehension by the students.

The use of Army TEAMThink across three different schools demonstrated that the program is not content specific and can be used in various content areas. Participants from the Engineer School, the Ordnance Munitions and Electronic Maintenance School, and the Military Intelligence School used Army TEAMThink, and demonstrated no substantial differences in performance across schools or content areas. This suggests that the principles behind Army TEAMThink would also be applicable to subjects beyond a military scope.



## References

Abell, M. (2000, December). Soldiers as distance learners: what army trainers need to know. Paper presented at the meeting at the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Orlando, FL. Retrieved April 25, 2002 from <http://www.tadlp.monroe.army.mil/abell%20paper.htm>.

Alkhateeb, H. M., & Jumaa, M. (2002). Cooperative learning and algebra performance of eighth grade students in united arab emirates. *Psychological Reports*, 90, 91-100.

Alliger, G.M., Tannenbaum, S.I., Bennett, W., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, 50, 341-358.

American Psychological Association. (1997). Learner-centered psychological principles: A framework for school redesign and reform. Washington, DC: Author. Retrieved May 7, 2002 from <http://www.apa.org/ed/lcp.html>.

Blanchette, J. (2001). Questions in the online learning environment. *Journal of Distance Education/Revue de l'enseignement à distance*: 16,2. Retrieved February 22, 2002 from <http://cade.icaap.org/vol16.2/blanchette.html>.

Bonk, C. J., & Kim, K. A. (1998). Extending sociocultural to adult learning. In M. C. Smith & T. Pourchot (Ed.), *Adult learning and development: Perspectives from educational psychology* (pp. 67-88). Mahwah, NJ: Lawrence Erlbaum Associates.

Bonk, C. J., & Wisher, R. A. (2000). Applying collaborative and e-learning tools to military distance-learning: A research framework. (Technical Report 1107). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.

Deatz, R. C., & Campbell, C. H. (2001). Application of cognitive principles in distributed computer-based training (Research Product 2001-03). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Defense Manpower Data Center (2002). Department of defense military manpower training report, FY2002. Author.

Duffy, T. M., Dueber, B., & Hawley, C. L. (1998). Critical thinking in a distributed environment: a pedagogical base for the design of conferencing systems. In C. J. Bonk & K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (pp. 51-78). Mahwah, NJ: Erlbaum..

Fisher, M., & Coleman, B. (2001-2002). Collaborative online learning in virtual discussions. *Journal of Educational Technology Systems*, 30(1), 3-17.

Freeman, M., Wisher, R. A., Curnow, C., & Morris, K. (1999). Down the digital dirt roads: Increasing access to distance-learning with hybrid audiographics. *Proceedings of the 1999 Interservice/Industry Training Simulations and Education Conference*, Arlington, VA: National Training Systems Association, 52-61.

Garrison, D. R., Anderson, T. and Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *The American Journal of Distance Education*, 15(1), 7-23.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, 31 (1), 104-137.

Graesser, A. C., Person, N., & Huber, J. (1992). Mechanisms that generate questions. In T. W. Lauer, E. Peacock, & A. C. Graesser (Eds.), *Questions and information systems* (pp.167-187). Hillsdale, NJ: L. Erlbaum.

Graesser, A. C., & Wisher, R. A. (2001). Question generation as a learning multiplier in distributed-learning environments (Technical Report 1121). Alexandria, VA; U.S. Army Research Institute for the Social and Behavioral Sciences.

Howard, F. S. (1992). Distributed training pilot – MOS 75B distributed training course training effectiveness analysis (TRAC-WSMR-TEA-92-031). US Army TRADOC Analysis Center-White Sands Missile Range.

Kang, I. (1998). The use of computer-mediated communication: Electronic collaboration and interactivity. In C. J. Bonk & K. S. King (Eds.), *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse* (pp. 315-337). Mahwah, NJ: Erlbaum..

Metzke, J., Redding, G. A., & Fletcher, J. D. (1996). Distance-learning and the reserve components. Alexandria, VA: Institute for Defense Analysis.

Minstrell, J. (1999). Expertise in teaching. In R. J. Sternberg & J.A. Horvath (Eds.) *Tacit knowledge in professional practice: Researcher and practitioner perspectives* (pp.215-230). Mahwah, NJ: Lawrence Erlbaum Associates.

Miyake, N., & Norman, D. A. (1979). To ask a question, one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, 18, 357-364.

Moore, M. G. (2001) Surviving as a distance teacher. *The American Journal of Distance Education*, 15(2), 1-5.

Moore, M. G. (1989). Three types of interaction. *The American Journal of Distance Education*, 3 (2), 1-6.

Muilenburg, L., & Berge, Z. L. (2001). Barriers to distance education: A factor-analytic study. *The American Journal of Distance Education*, 15(2), 7-24.

Navarro, P., & Shoemaker, J. (2000). Performance and perceptions of distance learners in cyberspace. *The American Journal of Distance Education*, 14(2), 15-35.

Otero, J., & Graesser, A.C. (2001). PREG: Elements of a model of question asking. *Cognition & Instruction*, 19, 143-175.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49, 345-75.

Program Management Office, The Army Distance-learning Program (2000). *The Army Distance-learning Program Brochure*, Fort Monroe, VA: Department of the Army. Retrieved from <http://www.tadlp.army.mil/brochure.pdf>

Savery, J. R. (1998). Fostering ownership with computer supported collaborative writing in higher education. In C. J. Bonk & K. S. King (Eds.) *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse*. Mahwah, NJ: Erlbaum..

Shavelson, R. J., Berliner, D. C., Ravitch, M. M., and Loeding, D. (1974). Effects of position and type of question on learning from prose material: Interaction of treatments with individual differences. *Journal of Educational Psychology*, 66, 40-48.

Swartz, M. L. (1987). Questions asked during command language learning: Implications for knowledge representation. (Technical Report 740). Alexandria, VA; U.S. Army Research Institute for the Social and Behavioral Sciences.

TRADOC (1999). *The Army Distance-learning Plan*. U.S. Army Training and Doctrine Command, Fort Monroe, VA.

Wisher, R.A. & Olson, T.M. (in review) The effectiveness of web-based instruction (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.



## Appendix A

### Examples of questions and comments from all three locations.

#### **Engineer School**

##### **Question:**

When developing a Scheme of Engineer Operations (SOEO) for use in the defense, which of the following criteria should you include for an obstacle group? (As per the question author, the correct answer is "d")

- a) Task and Purpose
- b) Relative location and the maneuver unit it supports
- c) Indirect fires allocated to an obstacle group
- d) All of the above

##### **Comments:**

- 1) Good Question
- 2) Unsure if obstacles have task and purpose. Obstacles have intent, target, and location.  
Units have  
task and purpose not obstacles. Revise and rewrite.
- 3) Task and purpose?????
- 4) Good Question

#### **OM&M School**

##### **Question:**

When transporting radioactive material, the proper label to affix to the package is based on what radiation level reading? (As per question author, correct answer is "a")

- a) The radiation level at the surface of the container
- b) The radiation level of the materials being transported
- c) The radiation level of the carrier from 100ft away
- d) The radiation level inside the sealed carrier

##### **Comments:**

- 1) Trick question, huh?
- 2) Suggestion, to make your question clearer, perhaps you should re-word it to read,  
"When transporting radioactive material the proper label that is affixed to the package is  
based upon..."
- 3) Could everyone with an answer please honestly tell me if they researched the question  
or came up with an educated guess?

## **Military Intelligence School**

Question:

How can we predict a terrorist act? (As per question author, correct answer is "b")

- a) more explosives being bought
- b) tip off from a known source
- c) previously unknown people asking specific questions about a compound or a building
- d) call 1-800 call mad bombers and ask for a reference

Comments:

- 1) The answers to this question do not make sense because the first three answers are all correct.
- 2) Can be misleading since the first three answer can be indicator...may want to rephrase question and ask which is more reliable/valid
- 3) Choices A, B, and C are all good answers
- 4) Both are indicators that would logically indicate that sort of or similar activity might occur
- 5) Prove that answer C is not indicator
- 6) No comment